

# Particle dynamics inside shocks in Hamilton–Jacobi equations

Kostya Khanin<sup>1</sup> and Andrei Sobolevski<sup>2,3</sup>

<sup>1</sup> Department of Mathematics, University of Toronto, Toronto, Ontario, Canada

<sup>2</sup> Institute for Information Transmission Problems of the Russian Academy of Sciences, Moscow, Russia

<sup>3</sup> Laboratoire J.-V. Poncelet (UMI 2615 CNRS), Moscow, Russia

E-mail: [khanin@math.toronto.edu](mailto:khanin@math.toronto.edu), [sobolevski@iitp.ru](mailto:sobolevski@iitp.ru)

**Abstract.** Characteristic curves of a Hamilton–Jacobi equation can be seen as action minimizing trajectories of fluid particles. For nonsmooth “viscosity” solutions, which give rise to discontinuous velocity fields, this description is usually pursued only up to the moment when trajectories hit a shock and cease to minimize the Lagrangian action. In this paper we show that for any convex Hamiltonian there exists a uniquely defined canonical global nonsmooth coalescing flow that extends particle trajectories and determines dynamics inside the shocks. We also provide a variational description of the corresponding effective velocity field inside shocks, and discuss relation to the “dissipative anomaly” in the limit of vanishing viscosity.

PACS numbers: 47.10.Df, 47.40.Nm, 02.30.Yy, 02.40.Ft, 02.40.Xx

Submitted to: *Philosophical Transactions of the Royal Society A*

## 1. Introduction

The Hamilton–Jacobi equation

$$\frac{\partial \phi}{\partial t}(t, \mathbf{x}) + H(t, \mathbf{x}, \nabla \phi(t, \mathbf{x})) = 0 \quad (1)$$

plays an important role in a large variety of mathematical and physical problems. Apart from analytical mechanics, it appears in description of a whole range of extended dissipative systems featuring nonequilibrium turbulent processes, from microscales of condensed matter and statistical physics through mesoscale setting of free-boundary fluid to macroscale cosmological evolution (see, e.g., a non-exhaustive collection of references in [1]). The central issue in a study of nonlinear evolution for (1) is to understand, both from the mathematical and physical points of view, the behaviour of the system after the inevitable formation of singularities.

A theory of weak solutions for a general Hamilton–Jacobi equation, employing the regularization by infinitesimal diffusion, exists since the 1970s [2, 3, 4]. In the one-dimensional setting this theory is essentially equivalent to the earlier theory of hyperbolic conservation laws in fluid mechanics, developed in the 1950s [5, 6, 7]. In more than one dimension, however, the two theories are no longer parallel.

The theory of weak solutions for the Hamilton–Jacobi equation is closely related to calculus of variations, and from this point of view one can say that introduction of diffusion is motivated essentially by stochastic control arguments [8]. In the present paper we adopt a related but somewhat complementary viewpoint, in which the Hamilton–Jacobi equation is considered as a fluid dynamics model, and construct the flow of “fluid particles” inside the shock singularities of a weak solution.

A useful example to be borne in mind when thinking about (1) — and arguably the most widely known variant thereof — is the Riemann, or inviscid Burgers, equation

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = 0, \quad \nabla \times \mathbf{u} = 0. \quad (2)$$

It is obtained for the Hamiltonian  $H(t, x, \mathbf{p}) = |\mathbf{p}|^2/2$  by setting  $\mathbf{u}(t, \mathbf{x}) = \nabla \phi(t, \mathbf{x})$ . Note that in (2) it is essential that the velocity field  $\mathbf{u}$  is curl-free, so this model is indeed equivalent to the Hamilton–Jacobi equation (1). The Riemann equation may in turn be considered as a limit of vanishing viscosity of the Burgers equation

$$\frac{\partial \mathbf{u}^\mu}{\partial t} + \mathbf{u}^\mu \cdot \nabla \mathbf{u}^\mu = \mu \nabla^2 \mathbf{u}^\mu, \quad \nabla \times \mathbf{u}^\mu = 0, \quad (3)$$

so solutions of (2) can be defined as limits of smooth solutions to (3) as the positive parameter  $\mu$  goes to zero.

The Burgers equation is in fact very special: it can be exactly mapped by the Cole–Hopf transformation into the linear heat equation and therefore explicitly integrated [5]. Nonetheless the qualitative behaviour of solutions to a parabolic regularization of (1) for a general convex Hamiltonian

$$\frac{\partial \phi^\mu}{\partial t} + H(t, \mathbf{x}, \nabla \phi^\mu) = \mu \nabla^2 \phi^\mu \quad (4)$$

in the limit of vanishing viscosity is similar to that for the Burgers equation. It turns out that as  $\mu \rightarrow 0$  there exists a limit  $\phi(t, \mathbf{x}) = \lim \phi^\mu(t, \mathbf{x})$  which is called the *viscosity solution*. Remarkably the viscosity solution can be described by a purely variational construction which does not use the viscous regularization at all. Below we briefly recall the main ideas of this variational approach.

Assume that the Hamiltonian function  $H(t, \mathbf{x}, \mathbf{p})$  is smooth and strictly convex in the momentum variable  $\mathbf{p}$ , i.e., is such that for all  $(t, \mathbf{x})$  the graph of  $H(t, \mathbf{x}, \mathbf{p})$  as a function of  $\mathbf{p}$  lies above any tangent plane and contains no straight segments. This implies that the velocity  $\mathbf{v} = \nabla_{\mathbf{p}} H(t, \mathbf{x}, \mathbf{p})$  is a one-to-one function of  $\mathbf{p}$ . In addition, the Lagrangian function

$$L(t, \mathbf{x}, \mathbf{v}) = \max_{\mathbf{p}} [\mathbf{p} \cdot \mathbf{v} - H(t, \mathbf{x}, \mathbf{p})]$$

under the above hypotheses is smooth and strictly convex in  $\mathbf{v}$ , although it may not be everywhere finite: e.g., the relativistic Hamiltonian  $H(t, \mathbf{x}, \mathbf{p}) = \sqrt{1 + |\mathbf{p}|^2}$  corresponds to the Lagrangian  $L(t, \mathbf{x}, \mathbf{v})$  that is defined for  $|\mathbf{v}| \leq 1$  as  $-\sqrt{1 - |\mathbf{v}|^2}$  and should be considered as taking value  $+\infty$  elsewhere. This will not happen if in addition one assumes that the Hamiltonian  $H$  grows superlinearly in  $|\mathbf{p}|$ .

The relation between the Lagrangian and the Hamiltonian is symmetric: they are Legendre conjugate to one another. This relation can also be expressed in the form of the Young inequality:

$$L(t, \mathbf{x}, \mathbf{v}) + H(t, \mathbf{x}, \mathbf{p}) \geq \mathbf{p} \cdot \mathbf{v}.$$

This inequality holds for all  $\mathbf{v}$  and  $\mathbf{p}$  and turns into equality whenever  $\mathbf{v} = \nabla_{\mathbf{p}} H(t, \mathbf{x}, \mathbf{p})$  or  $\mathbf{p} = \nabla_{\mathbf{v}} L(t, \mathbf{x}, \mathbf{v})$ . The two functions  $\nabla_{\mathbf{p}} H(t, \mathbf{x}, \mathbf{p})$  and  $\nabla_{\mathbf{v}} L(t, \mathbf{x}, \mathbf{v})$  are inverse to each other; we will call them the *Legendre transforms* at  $(t, \mathbf{x})$  of  $\mathbf{p}$  and of  $\mathbf{v}$ . (Usually the term “Legendre transform” refers to the relation between the conjugate functions  $H$  and  $L$ ; here we follow the usage adopted by A. Fathi in his works on weak KAM theory [9], which is more convenient in the present context.)

Note that if  $H(t, \mathbf{x}, \mathbf{p}) = |\mathbf{p}|^2/2$ , then  $L(t, \mathbf{x}, \mathbf{v}) = |\mathbf{v}|^2/2$  and the Legendre transform reduces to the identity  $\mathbf{v} \equiv \mathbf{p}$ , blurring the distinction between velocities and momenta. This is another very special feature of the Burgers equation.

Now assume that  $\phi(t, \mathbf{x})$  is a strong solution of the inviscid equation (1), i.e., a smooth function that satisfies the equation in the classical sense. For an arbitrary trajectory  $\gamma(t)$  the full time derivative of  $\phi$  along  $\gamma$  is given by

$$\frac{d\phi(t, \gamma)}{dt} = \frac{\partial \phi}{\partial t} + \nabla \phi \cdot \dot{\gamma} = \nabla \phi \cdot \dot{\gamma} - H(t, \gamma, \nabla \phi) \leq L(t, \gamma, \dot{\gamma}), \quad (5)$$

where at the last step the Young inequality is used. This implies a bound for the mechanical action corresponding to the trajectory  $\gamma$ :

$$\phi(t_2, \gamma(t_2)) \leq \phi(t_1, \gamma(t_1)) + \int_{t_1}^{t_2} L(s, \gamma(s), \dot{\gamma}(s)) ds. \quad (6)$$

Equality in (5) is only achieved if  $\dot{\gamma}$  is the Legendre transform of  $\nabla \phi$  at  $(t, \gamma)$ :

$$\dot{\gamma}(t) = \nabla_{\mathbf{p}} H(t, \gamma, \nabla \phi(t, \gamma)). \quad (7)$$

This represents one of Hamilton’s canonical equations, with momentum given for the trajectory  $\gamma$  by  $\mathbf{p}_\gamma(t) := \nabla\phi(t, \gamma(t))$ . The other canonical equation,  $\dot{\mathbf{p}} = -\nabla_{\mathbf{x}}H$ , follows from (1) and (7) because

$$\dot{\mathbf{p}}_\gamma(t) = \frac{\partial \nabla\phi}{\partial t} + (\nabla \otimes \nabla\phi) \cdot \dot{\gamma} = -\nabla_{\mathbf{x}}H(t, \gamma, \nabla\phi) - \nabla_{\mathbf{p}}H \cdot (\nabla \otimes \nabla\phi) + (\nabla \otimes \nabla\phi) \cdot \dot{\gamma}$$

Therefore the bound (6) is achieved for trajectories satisfying Hamilton’s canonical equations. This is a manifestation of the variational *principle of the least action*: Hamiltonian trajectories  $(\gamma(t), \mathbf{p}_\gamma(t))$  are (locally) action minimizing. In particular, if the initial condition

$$\phi(t=0, \mathbf{y}) = \phi_0(\mathbf{y}), \quad (8)$$

is a fixed smooth function, the identity

$$\phi(t, \mathbf{x}) = \phi_0(\gamma(0)) + \int_0^t L(s, \gamma(s), \dot{\gamma}(s)) ds$$

holds for a minimizer  $\gamma$  such that  $\gamma(t) = \mathbf{x}$ .

The least action principle can be used to *construct* the viscosity solution corresponding to the initial data (8):

$$\phi(t, \mathbf{x}) = \min_{\gamma: \gamma(t)=\mathbf{x}} \left( \phi_0(\gamma(0)) + \int_0^t L(s, \gamma(s), \dot{\gamma}(s)) ds \right). \quad (9)$$

This is the celebrated Lax–Oleinik formula (see, e.g., [10] or [9]), which reduces a PDE problem (1), (8) to the variational problem (9) where minimization is extended to all smooth curves  $\gamma$  such that  $\gamma(t) = \mathbf{x}$ .

The function  $\phi$  defined by (9) is smooth at those points  $(t, \mathbf{x})$  where the minimizing trajectory is unique. In this case, the minimizer can be embedded in a smooth family of minimizing trajectories whose endpoints at time 0 and  $t$  are continuously distributed about  $\gamma(0)$  and  $\gamma(t) = \mathbf{x}$ . A piece of initial data  $\phi_0$  gets continuously deformed according to (5) along this bundle of trajectories into a piece of smooth solution  $\phi$  to (1) defined in a neighbourhood of  $\mathbf{x}$  at time  $t$ . Of course the Hamilton–Jacobi equation is satisfied in a strong sense at all points where  $\phi$  is differentiable.

However, the crucial feature of (9) is that generally there will be points  $(t, \mathbf{x})$  with several minimizers  $\gamma_i$ , which start at different locations  $\gamma_i(0)$  but bring the same value of action to  $\mathbf{x} = \gamma_i(t)$ . Just as above, each of these Hamiltonian trajectories will be responsible for a separate smooth “piece” of solution. Thus for locations  $\mathbf{x}'$  close to  $\mathbf{x}$  the function  $\phi$  will be represented as a pointwise minimum of smooth pieces  $\phi_i$ :

$$\phi(t, \mathbf{x}') = \min_i \phi_i(t, \mathbf{x}').$$

As all  $\gamma_i$  have the same terminal value of action, all these pieces intersect at  $(t, \mathbf{x})$ :  $\phi_1(t, \mathbf{x}) = \phi_2(t, \mathbf{x}) = \dots = \phi(t, \mathbf{x})$ . Thus the neighbourhood of  $\mathbf{x}$  at time  $t$  is partitioned into domains where  $\phi$  coincides with each of the smooth functions  $\phi_i$  and satisfies the Hamilton–Jacobi equation (1) strongly. These domains are separated by surfaces of various dimensions where two, or possibly three or more, pieces  $\phi_i$  intersect and hence  $\phi$  is not differentiable. Such surfaces are called *shocks*. Note that

a function  $\phi$  defined by the Lax–Oleinik formula is continuous everywhere, including the shocks; it is its gradient that suffers a discontinuity. In general, there are infinitely many continuous functions that match the initial condition (8) and are differentiable and satisfy the Hamilton–Jacobi equation (1) apart from some shock surfaces, just as  $\phi$ . A standard one-dimensional example of such nonuniqueness is provided by  $\phi_\alpha(t, x) = \min(\alpha|x| - \alpha^2 t/2, 0)$ , which for any  $\alpha > 0$  satisfies the initial condition  $\phi_\alpha(0, x) = 0$  and the equation  $\partial\phi_\alpha/\partial t + |\partial\phi_\alpha/\partial x|^2/2 = 0$  apart from the shock rays  $x = \pm\alpha t/2, x = 0$ . What distinguishes the function  $\phi$  defined by (9) from all these “weak solutions,” and grants it with important physical meaning, is that  $\phi$  appears in the limit of vanishing viscosity for the regularized equation (4) with the initial condition (8) (see, e.g., [3]).

For a smooth Hamiltonian it can be proved that once shocks are created they never disappear, although they can merge with one another. Another important physical feature of viscosity solutions is that minimizers can only merge with shocks but never leave them: all minimizers coming to some  $(t, \mathbf{x})$  in (9) originate at  $t = 0$ . It is easy to see that this is not so in the above example of  $\phi_\alpha$ , where minimizers emerge from  $x = 0$  at all times  $t > 0$ .

Moreover, in a solution  $\phi$  given by (9) a minimizer that has come to a shock cannot be continued any longer as a minimizing trajectory. Indeed, wherever it comes, there will be other trajectories originated at  $t = 0$  that will bring smaller values of action to the same location. Hence for the purpose of the least action description (9) Hamiltonian trajectories are discontinued as soon as they are absorbed by shocks. The set of trajectories which survive as minimizers until time  $t > 0$  is decreasing with  $t$ , but at all times it is sufficiently large to cover the whole space of final positions.

Let us now adopt an alternative viewpoint and consider the Hamilton–Jacobi equation as a fluid dynamics model, assuming that Hamiltonian trajectories (7) are described by material “particles” transported by the velocity field  $\mathbf{u}(t, \mathbf{x})$ , which is the Legendre transform of the momenta field  $\mathbf{p}(t, \mathbf{x}) = \nabla\phi(t, \mathbf{x})$ . From this new perspective it is no longer natural to accept that particles annihilate once they reach a shock. Can therefore something be said about the dynamics of those particles that got into the shock, notwithstanding the fact that their trajectories cease to minimize the action? The problem here comes from the discontinuous nature of the velocity field  $\mathbf{u}$ , which makes it impossible to construct classical solutions to the transport equation  $\dot{\gamma}(t) = \mathbf{u}(t, \gamma)$ .

In dimension 1 the answer to the question above is readily available. Shocks at each fixed  $t$  are isolated points in the  $\mathbf{x}$  space and as soon as a trajectory merges with one of them, it continues to move with the shock at all later times. This description is related to C. Dafermos’ theory of generalized characteristics [11] which, in fact, can be extended to a much more general situation of nonconvex Hamiltonians and systems of conservation laws. However, in several space dimensions shocks become extended surfaces and already for equation (1) with a strictly convex Hamiltonian dynamics of trajectories inside shocks is by no means trivial. The main goal of the present paper is to describe a natural and canonical construction of such dynamics.

First results in this direction were obtained by I. Bogaevsky [12, 13] for the Burgers equation (3). Bogaevsky suggested to consider the transport problem for a smooth velocity field  $\mathbf{u}^\mu$ :

$$\dot{\gamma}^\mu(t) = \mathbf{u}^\mu(t, \gamma^\mu), \quad \gamma^\mu(0) = \mathbf{y}.$$

Since  $\mathbf{u}^\mu$  for  $\mu > 0$  is a smooth vector field, this equation defines a family of particle trajectories forming a smooth flow. The next step is to take the limit of this flow as  $\mu \downarrow 0$ . Bogaevsky proved that this limit exists as a non-differentiable continuous flow, for which the forward derivative  $\dot{\gamma}(t+0) = \lim_{\tau \downarrow 0} [\gamma(t+\tau) - \gamma(t)]/\tau$  is defined everywhere. If  $\gamma(t)$  is located outside shocks, this derivative coincides with  $\mathbf{u}(t, \gamma(t))$ . Otherwise there are several limit values of velocity  $\mathbf{u}_i$ , and Bogaevsky discovered an interesting explicit representation for  $\dot{\gamma}(t+0)$ : it coincides with the center of the smallest ball that contains all  $\mathbf{u}_i$ . It should be remarked that uniqueness of a limit flow in the case of a quadratic Hamiltonian was earlier observed by P. Cannarsa and C. Sinestrari in the context of propagation of singularities for the eikonal equation and differential inclusions [14, Lemma 5.6.2].

The original approach in [12, 13] is based on the specific properties of the Burgers equation and cannot be applied in the case of general convex Hamiltonians. In particular, the method uses the identity of velocities and momenta, which does not hold in the general setting. At the same time the common wisdom says that all Hamilton–Jacobi equations with convex Hamiltonians must have similar properties.

In this work we consider the above strategy, consisting in the parabolic regularization of the Hamilton–Jacobi equation and investigation of the vanishing viscosity limit for the corresponding regularized flow, in the case of general convex Hamiltonians. We show that such a limit exists, and derive an explicit representation for the forward velocity of the limit flow that extends the above result for the Burgers equation. Yet the mechanism powering these results in the general case is completely different. It is based on the fundamental uniqueness of a possible limit behaviour of  $\gamma^\mu$ , which we discuss in detail below.

The paper is organized as follows. In Section 2 we study the local structure of a viscosity solution near a singularity. We also introduce the notion of admissible velocity at a singularity and show that it can be determined uniquely. Moreover, the unique admissible velocity provides a solution to a certain convex minimization problem, which generalizes Bogaevsky’s construction of the center of the smallest ball. In Section 3 we demonstrate that for any nonsmooth viscosity solution there exists a unique continuous nonsmooth flow of trajectories tangent to the field of admissible velocities. Section 4 contains concluding remarks.

## Acknowledgments

This work is supported by the NSERC Discover Grant and the Russian Foundation for Basic Research, project RFBR-CNRS–07–01–92217; the second author is supported in

part by the French Agence Nationale de la Recherche, project ANR–07–BLAN–0235 OTARIE. We acknowledge the hospitality of Observatoire de la Côte d’Azur (Nice, France), where part of this work has been performed, as well as support of the French Ministry for National Education.

We are sincerely grateful to Jérémie Bec, Patrick Bernard, Ilya Bogaevsky, Yann Brenier, Philippe Choquard, and Boris Khesin for numerous valuable discussions. It is our special pleasure to express gratitude to Uriel Frisch, who made crucial contributions to the current revival of interest in the “Burgers turbulence” and related fields, and who is turning 70 in 2010.

## 2. Local structure of viscosity solutions and admissible momenta

Let  $\phi$  be a viscosity solution to the Hamilton–Jacobi equation (1) with initial data (8). If there is a single minimizer coming to  $(t, \mathbf{x})$ , then  $\phi$  is differentiable at this point and

$$\begin{aligned}\phi(t + \tau, \mathbf{x} + \boldsymbol{\xi}) &= \phi(t, \mathbf{x}) + \frac{\partial \phi}{\partial t} \tau + \nabla \phi \cdot \boldsymbol{\xi} + \dots \\ &= \phi(t, \mathbf{x}) - H(t, \mathbf{x}, \nabla \phi) \tau + \nabla \phi \cdot \boldsymbol{\xi} + \dots,\end{aligned}$$

where dots  $\dots$  stand for higher-order terms. If  $\phi$  is not differentiable at  $(t, \mathbf{x})$ , this means that there are several minimizers  $\gamma_i$  such that  $\gamma_i(t) = \mathbf{x}$ , each bringing to  $(t, \mathbf{x})$  a different piece  $\phi_i$  of solution. Then the Lax–Oleinik formula implies that

$$\begin{aligned}\phi(t + \tau, \mathbf{x} + \boldsymbol{\xi}) &= \min_i \phi_i(t + \tau, \mathbf{x} + \boldsymbol{\xi}) \\ &= \phi(t, \mathbf{x}) + \min_i (-H_i \tau + \mathbf{p}_i \cdot \boldsymbol{\xi}) + \dots,\end{aligned}$$

where  $\mathbf{p}_i := \nabla \phi_i(t, \mathbf{x})$  and  $H_i := H(t, \mathbf{x}, \mathbf{p}_i)$ .

In the latter case neither of the expressions  $-H_i \tau + \mathbf{p}_i \cdot \boldsymbol{\xi}$  provides an adequate linear approximation to the difference  $\phi(t + \tau, \mathbf{x} + \boldsymbol{\xi}) - \phi(t, \mathbf{x})$ , but they all *majorize* this difference up to a remainder that is linear or higher-order depending on  $\tau$  and  $\boldsymbol{\xi}$ . Evidently, so does too the linear form  $-H \tau + \mathbf{p} \cdot \boldsymbol{\xi}$  for any convex combination

$$\mathbf{p} = \sum_i \lambda_i \mathbf{p}_i, \quad H = \sum_i \lambda_i H_i$$

with  $\lambda_i \geq 0$ ,  $\sum_i \lambda_i = 1$ . In convex analysis these convex combinations are called supergradients of  $\phi$  at  $(t, \mathbf{x})$  and the whole collection of them, which is a convex polytope with vertices  $(-H_i, \mathbf{p}_i)$ , is called the *superdifferential* of  $\phi$  [14, 15]. We use R. T. Rockafellar’s notation  $\partial \phi(t, \mathbf{x})$  for the superdifferential [15].

To avoid a possible misunderstanding it should be noted that, although uniqueness of minimizer coming to  $(t, \mathbf{x})$  implies differentiability of  $\phi$  at  $t$  and earlier times, it does *not* imply its differentiability at any  $t + \tau > t$ . The following example shows how this may happen. The function defined for  $\tau \geq 0$  by

$$\phi(t + \tau, x + \xi) = \phi(t, x) - \frac{4}{3} \tau^3 - 2|\xi| \tau - \frac{4}{3} (|\xi| + \tau^2)^{3/2}$$

is a viscosity solution of the equation  $\partial \phi / \partial t + |\partial \phi / \partial x|^2 / 2 = 0$  that satisfies the smooth initial condition  $\phi(t, x + \xi) - \phi(t, x) = -\frac{4}{3} |\xi|^{3/2}$ . For  $\tau > 0$  a shock appears at  $\xi = 0$ , but

differentiability at  $\tau = 0$  is recovered because  $\partial\phi(t + \tau, x) = \{-8\tau^2\} \times [-4\tau, 4\tau]$  shrinks to  $(0, 0)$  as  $\tau \downarrow 0$ . Such points  $(t, \mathbf{x})$  are called *preshocks* [1].

The particular case of preshocks is an instance of a general fact: replacing gradients with superdifferentials allows to recover continuous differentiability, but in a weaker sense. Namely, suppose  $(t_n, \mathbf{x}_n)$  converges to  $(t, \mathbf{x})$  and the sequence  $(-H_n, \mathbf{p}_n) \in \partial\phi(t_n, \mathbf{x}_n)$  has a limit point  $(-H, \mathbf{p})$ . By definition of superdifferential,

$$\phi(t_n + \tau, \mathbf{x}_n + \boldsymbol{\xi}) - \phi(t_n, \mathbf{x}_n) \leq -H_n \tau + \mathbf{p}_n \cdot \boldsymbol{\xi} + \cdots; \quad (10)$$

passing here to the limit and using continuity of  $\phi$ , we see that  $(-H, \mathbf{p}) \in \partial\phi(t, \mathbf{x})$ . Therefore the superdifferential  $\partial\phi(t, \mathbf{x})$  contains all the limit points of superdifferentials  $\partial\phi(t_n, \mathbf{x}_n)$  as  $(t_n, \mathbf{x}_n)$  converges to  $(t, \mathbf{x})$ .

Suppose  $(t, \mathbf{x})$  is a point of shock where  $k$  smooth branches  $\phi_i$  meet. It follows from the above discussion that for a particle moving from a shock point  $(t, \mathbf{x})$  all possible values of the velocity  $\mathbf{v}$  must be such that the corresponding momentum  $\mathbf{p}(\mathbf{v})$  belongs to the convex hull of the available momenta  $\mathbf{p}_i = \nabla\phi_i(t, \mathbf{x})$ ,  $1 \leq i \leq k$ . However, one can say even more. For small positive  $\tau$  not all the branches  $\phi_i$  will contribute to the solution  $\phi$  at a point  $(t + \tau, x + \mathbf{v}\tau)$ , but only those of them that correspond to a minimum in  $\min_i(-H_i + \mathbf{p}_i \cdot \mathbf{v})$ . Denote the corresponding set of indices

$$I(\mathbf{v}) := \{1 \leq j \leq k: -H_j + \mathbf{p}_j \cdot \mathbf{v} = \min_i(-H_i + \mathbf{p}_i \cdot \mathbf{v})\}. \quad (11)$$

The set  $I(\mathbf{v})$  has the following physical meaning: if particle moves away from a shock with a given velocity  $\mathbf{v}$  then only  $\phi_j$  and  $\mathbf{p}_j$  for  $j \in I(\mathbf{v})$  are relevant. Geometrically one can say that the convex hull of  $\mathbf{p}_j$ ,  $j \in I(\mathbf{v})$ , is the  $\mathbf{p}$ -projection of the face of the superdifferential  $\partial\phi(t, \mathbf{x})$  that looks toward an infinitesimal observer who has just left  $(t, \mathbf{x})$  with velocity  $\mathbf{v}$ .

This implies that any possible velocity  $\mathbf{v}$  must satisfy the following condition.

**Admissibility Condition.** *A velocity  $\mathbf{v}^*$  is admissible if and only if the corresponding momentum belongs to the convex hull of the momenta  $\mathbf{p}_j$ ,  $j \in I(\mathbf{v}^*)$ . Namely,*

$$\mathbf{p}^*(\mathbf{v}^*) = \sum_{j \in I(\mathbf{v}^*)} \lambda_j \mathbf{p}_j, \quad \lambda_j \geq 0, \quad \sum_{j \in I(\mathbf{v}^*)} \lambda_j = 1. \quad (12)$$

To make the admissibility argument rigorous it is necessary to have some control over the remainder term in (10). A natural function class that contains viscosity solutions of Hamilton–Jacobi equations and in which such control is possible is formed by semiconcave functions [14]. We refer a reader interested in careful proofs of this and other convex analytic results used in this paper to monographs [14, 15].

Remarkably the Admissibility Condition determines the velocity  $\mathbf{v}^*$  uniquely.

**Lemma 1 (Uniqueness)** *Let  $\phi$  be a viscosity solution to the Cauchy problem (1), (8). Then at any  $(t, \mathbf{x})$  there exists a unique admissible velocity  $\mathbf{v}^*$ . Moreover this admissible velocity  $\mathbf{v}^*$  is the unique point of the global minimum for the following function*

$$\hat{L}(\mathbf{v}) := L(t, \mathbf{x}, \mathbf{v}) - \min_i(-H_i + \mathbf{p}_i \cdot \mathbf{v}). \quad (13)$$



*Proof.* Recall that  $L(t, \mathbf{x}, \mathbf{v})$  is a strictly convex function of  $\mathbf{v}$  because of assumptions formulated in the Introduction. Rewriting

$$L_i(\mathbf{v}) := L(t, \mathbf{x}, \mathbf{v}) + H_i - \mathbf{p}_i \cdot \mathbf{v}, \quad \hat{L}(\mathbf{v}) = \max_i L_i(\mathbf{v}), \quad (14)$$

we see that  $\hat{L}(\mathbf{v})$  is a pointwise maximum of a finite number of strictly convex functions and therefore is strictly convex itself. Furthermore, because the Hamiltonian  $H(t, \mathbf{x}, \mathbf{p})$  is assumed to be finite for all  $\mathbf{p}$ , its conjugate Lagrangian  $L(t, \mathbf{x}, \mathbf{v})$  grows faster than any linear function as  $|\mathbf{v}|$  increases, and all its level sets are bounded. Thus  $\hat{L}(\mathbf{v})$  attains its minimum at a unique value of velocity  $\mathbf{v}^*$ .

Let us show that this point of minimum  $\mathbf{v}^*$  satisfies the admissibility condition. Indeed,

$$\nabla_{\mathbf{v}} L_i(\mathbf{v}^*) = \nabla_{\mathbf{v}} L(t, \mathbf{x}, \mathbf{v}^*) - \mathbf{p}_i = \mathbf{p}^* - \mathbf{p}_i.$$

Suppose that  $\mathbf{p}^*$  does not belong to the convex hull of  $\mathbf{p}_j$ ,  $j \in I(\mathbf{v}^*)$ . Then there exists a vector  $\mathbf{h}$  such that  $(\mathbf{p}^* - \mathbf{p}_j) \cdot \mathbf{h} < 0$  for all  $j \in I(\mathbf{v}^*)$ . It follows that  $L_j(\mathbf{v}^* + \epsilon \mathbf{h}) < L_j(\mathbf{v}^*)$  for all  $j \in I(\mathbf{v}^*)$  if  $\epsilon > 0$  is sufficiently small. Hence,  $\hat{L}(\mathbf{v}^* + \epsilon \mathbf{h}) < \hat{L}(\mathbf{v}^*)$  for sufficiently small  $\epsilon$ , which contradicts our assumption that  $\mathbf{v}^*$  is a point of minimum. This contradiction proves that  $\mathbf{v}^*$  is admissible.

To prove uniqueness we show that if  $\hat{\mathbf{v}}$  is admissible then it is a unique point of global minimum for the function  $\hat{L}$ . Using the strict convexity of  $\hat{L}$ , we obtain

$$\begin{aligned} L_j(\hat{\mathbf{v}} + \mathbf{h}) &= L(t, \mathbf{x}, \hat{\mathbf{v}} + \mathbf{h}) + H_j - \mathbf{p}_j \cdot (\hat{\mathbf{v}} + \mathbf{h}) \\ &> L_j(\hat{\mathbf{v}}) + \nabla_{\mathbf{v}} L(t, \mathbf{x}, \hat{\mathbf{v}}) \cdot \mathbf{h} - \mathbf{p}_j \cdot \mathbf{h} = L_j(\hat{\mathbf{v}}) + (\hat{\mathbf{p}} - \mathbf{p}_j) \cdot \mathbf{h}, \end{aligned}$$

where  $\hat{\mathbf{p}}$  is the Legendre transform of  $\hat{\mathbf{v}}$ . Since  $\hat{\mathbf{v}}$  is admissible,  $\hat{\mathbf{p}} = \sum_j \lambda_j \mathbf{p}_j$ , where all  $\lambda_j \geq 0$  and  $\sum_j \lambda_j = 1$ . Hence,  $\sum_j \lambda_j (\hat{\mathbf{p}} - \mathbf{p}_j) \cdot \mathbf{h} = [(\sum_j \lambda_j) \hat{\mathbf{p}} - \sum_j \lambda_j \mathbf{p}_j] \cdot \mathbf{h} = [\hat{\mathbf{p}} - \hat{\mathbf{p}}] \cdot \mathbf{h} = 0$ . It follows that  $(\hat{\mathbf{p}} - \mathbf{p}_j) \cdot \mathbf{h} \geq 0$  for at least one  $j \in I(\hat{\mathbf{v}})$ . Thus  $\hat{L}(\hat{\mathbf{v}} + \mathbf{h}) > \hat{L}(\hat{\mathbf{v}})$ , which implies that  $\hat{\mathbf{v}}$  is a unique point of global minimum for  $\hat{L}$ . This observation concludes the proof.

The admissibility property, first formulated above in a somewhat unmanageable combinatorial form (12), turns out to be the optimality condition in a convex minimization problem given by (13), i.e., a much simpler object. In particular, if  $\phi$  is differentiable at  $(t, \mathbf{x})$ , then  $\hat{L}(\mathbf{v}) = L(t, \mathbf{x}, \mathbf{v}) + H(t, \mathbf{x}, \nabla \phi) - \nabla \phi \cdot \mathbf{v}$  and the minimum in (13) is achieved at the Legendre transform of  $\nabla \phi$ . We thus recover Hamilton's equation (7).

The following reformulation will clarify the connection between admissibility and Bogaevsky's original construction for the Burgers equation. Let  $\mathbf{v}_i = \nabla_{\mathbf{p}} H(t, \mathbf{x}, \mathbf{p}_i)$  be the velocity corresponding to the limit momentum  $\mathbf{p}_i$  and observe that  $\mathbf{p}_i = \nabla_{\mathbf{v}} L(t, \mathbf{x}, \mathbf{v}_i)$ . The Young inequality implies that  $H_i = H(t, \mathbf{x}, \mathbf{p}_i) = \mathbf{p}_i \cdot \mathbf{v}_i - L(t, \mathbf{x}, \mathbf{v}_i)$  and therefore (14) assumes the form

$$\hat{L}(\mathbf{v}) = \max_i [L(t, \mathbf{x}, \mathbf{v}) - L(t, \mathbf{x}, \mathbf{v}_i) - \nabla_{\mathbf{v}} L(t, \mathbf{x}, \mathbf{v}_i) \cdot (\mathbf{v} - \mathbf{v}_i)].$$

The quantity in square brackets is known as the *Bregman divergence* between vectors  $\mathbf{v}$  and  $\mathbf{v}_i$ , a specific measure of their separation with respect to the convex function  $L$

[16]. When  $L(t, \mathbf{x}, \mathbf{v}) = |\mathbf{v}|^2/2$ , the Bregman divergence reduces to (half) the squared distance between the two vectors; hence the admissible velocity  $\mathbf{v}^*$  exactly coincides with the centre of smallest ball containing all  $\mathbf{v}_i$ , and Bogaevsky’s result is recovered.

Finally, we discuss the physical meaning of the function  $\hat{L}$ . Consider an infinitesimal movement from  $(t, \mathbf{x})$  with velocity  $\mathbf{v}$ . Obviously  $\phi(t, \mathbf{x}) + L(t, \mathbf{x}, \mathbf{v}) dt - \phi(t + dt, \mathbf{x} + \mathbf{v} dt) \geq 0$ . It is easy to see that in the linear approximation in  $dt$

$$\phi(t, \mathbf{x}) + L(t, \mathbf{x}, \mathbf{v}) dt - \phi(t + dt, \mathbf{x} + \mathbf{v} dt) = \hat{L}(\mathbf{v}) dt.$$

Hence the unique admissible velocity  $\mathbf{v}^*$  minimizes the rate of the difference in action between the true minimizers and trajectories of particles on shocks. In other words, the trajectory on a shock cannot be a minimizer but it does its best to keep its surplus action growing as slowly as possible.

### 3. The vanishing viscosity limit

In the preceding section we constructed a canonical vector field  $\mathbf{v}^* = \nabla_{\mathbf{p}} H(t, \mathbf{x}, \mathbf{p}^*)$  corresponding to a given viscosity solution  $\phi$  of the Cauchy problem (1), (8). The basis of this construction, the admissibility condition, appears as a natural consistency condition between velocities and supergradients. This condition together with the variational principle (13) guarantees uniqueness of the admissible pair  $(\mathbf{v}^*, \mathbf{p}^*)$ .

The vector field  $\mathbf{v}^*(t, \mathbf{x})$  can be decomposed into a union of smooth tangent vector fields defined on connected pieces of smooth shock surfaces of various dimensions as well as on the domain where  $\phi$  is differentiable, so dynamics in the latter domain or inside any piece of a smooth shock surface is given locally by a smooth flow. But globally the field  $\mathbf{v}^*$  is discontinuous, and it is not immediately clear if there exists an overall continuous flow of trajectories  $\gamma$  that is compatible with  $\mathbf{v}^*$  in the sense that  $\dot{\gamma}(t+0) = \mathbf{v}^*(t, \gamma)$ . Even less obvious is the uniqueness of such a flow.

In order to answer these questions affirmatively we employ the vanishing viscosity limit for the parabolic regularization

$$\frac{\partial \phi^\mu}{\partial t} + H(t, \mathbf{x}, \nabla \phi^\mu) = \mu \nabla^2 \phi^\mu, \quad \mu > 0, \quad (15)$$

of the Hamilton–Jacobi equation (1). For sufficiently smooth initial data  $\phi^\mu(t=0, \mathbf{y}) = \phi_0(\mathbf{y})$  equation (15) has a globally defined strong solution, which is locally Lipschitz with a constant independent of  $\mu$ . Moreover,  $\phi^\mu$  converges as  $\mu \downarrow 0$  to the unique viscosity solution  $\phi$  corresponding to the same initial data. Proof of these facts may be found, e.g., [3], where they are established for  $\phi_0 \in C^{2,\alpha}$ .

Consider now the differential equation

$$\dot{\gamma}^\mu(t) = \nabla_{\mathbf{p}} H(t, \gamma^\mu, \nabla \phi^\mu(t, \gamma^\mu)), \quad \gamma^\mu(0) = \mathbf{y}. \quad (16)$$

For  $\mu > 0$  this equation has a unique solution  $\gamma_y^\mu$ , which continuously depends on the initial location  $\mathbf{y}$ . Fix a point  $(t, \mathbf{x})$  with  $t > 0$  and pick for all sufficiently small  $\mu > 0$  trajectories  $\gamma^\mu$  such that  $\gamma^\mu(t) = \mathbf{x}$ . The uniform Lipschitz property of solutions  $\phi^\mu$  implies that the curves  $\gamma^\mu$  are uniformly bounded and equicontinuous on some interval

containing  $t$ . Hence there exists a curve  $\bar{\gamma}$  and a sequence  $\mu_i \downarrow 0$  such that  $\lim_{\mu_i \downarrow 0} \gamma^{\mu_i} = \bar{\gamma}$  uniformly. Note that all  $\gamma^{\mu_i}$  and  $\bar{\gamma}$  are also Lipschitz with a constant independent of  $\mu$  and that  $\bar{\gamma}(t) = \mathbf{x}$ . Let furthermore  $\bar{\mathbf{v}}$  be a limit point of the “forward velocity” of the curve  $\bar{\gamma}$  at  $(t, \mathbf{x})$ , i.e., let for some sequence  $\tau_k \downarrow 0$

$$\bar{\mathbf{v}} = \lim_{\tau_k \downarrow 0} \frac{1}{\tau_k} [\bar{\gamma}(t + \tau_k) - \bar{\gamma}(t)].$$

Of course neither the curve  $\bar{\gamma}$  nor the velocity  $\bar{\mathbf{v}}$  are *a priori* defined uniquely. Nevertheless it turns out that  $\bar{\mathbf{v}}$  must satisfy the admissibility condition with respect to the solution  $\phi$  and therefore it coincides with  $\mathbf{v}^*$ . Also, trajectories of the flow  $\gamma^\mu$  converge as  $\mu \downarrow 0$  to segments of integral trajectories of the vector field  $\mathbf{v}^*$  on smooth shock surfaces, establishing the uniqueness of the limit flow  $\bar{\gamma}$ . Moreover, the limit flow is *coalescing*: if two trajectories intersect at time  $t$ , they coincide for all  $t' > t$ . All these statements follow from the following fact.

**Lemma 2** *The flow defined by (16) for sufficiently small  $\mu > 0$  collapses the neighbourhood of the shock surface that passes through  $(t, \mathbf{x})$  and is tangent to  $(1, \mathbf{v}^*)$  onto this surface.*

Here is a sketch of proof. Let  $\mathbf{v}^*$  be the admissible velocity corresponding to  $(t, \mathbf{x})$ ,  $\mathbf{p}^*$  the corresponding momentum, and define

$$H^* = \mathbf{p}^* \cdot \mathbf{v}^* - \min_i (\mathbf{p}_i \cdot \mathbf{v}^* - H_i). \quad (17)$$

The full time derivative of the function  $\psi^\mu(t + \tau, \mathbf{x} + \boldsymbol{\xi}) = \phi^\mu(t + \tau, \mathbf{x} + \boldsymbol{\xi}) - \mathbf{p}^* \cdot \boldsymbol{\xi} + H^* \tau$  along an integral trajectory  $\gamma^\mu$  of equation (16) which passes through  $\mathbf{x} + \boldsymbol{\xi}$  at time  $t + \tau$  is given by

$$\frac{d\psi^\mu(t + \tau, \gamma^\mu(t + \tau))}{d\tau} = \frac{\partial \phi^\mu}{\partial t} + H^* + (\nabla \phi^\mu - \mathbf{p}^*) \cdot \dot{\gamma}^\mu.$$

Convergence of viscosity solutions implies convergence of their superdifferentials (to see this, it is enough to replace in (10) the function  $\phi$  with a sequence of functions  $\phi_n$  converging pointwise). Therefore limit points of  $(\partial \phi^\mu / \partial t, \nabla \phi^\mu)$  belong to  $\partial \phi(t, \mathbf{x})$  for all  $(t, \mathbf{x})$ , and for sufficiently small  $\mu > 0$ ,  $\tau > 0$ ,  $\boldsymbol{\xi}$  there exists  $(-H^\mu, \mathbf{p}^\mu) \in \partial \phi(t, \mathbf{x})$  and  $\mathbf{v}^\mu = \nabla_{\mathbf{p}} H(t, \mathbf{x}, \mathbf{p}^\mu)$  such that, to the linear approximation,  $\mathbf{v}^\mu = \dot{\gamma}^\mu + \dots$  and

$$\frac{d\psi^\mu}{d\tau} = -H^\mu + H^* + (\mathbf{p}^\mu - \mathbf{p}^*) \cdot \mathbf{v}^\mu + \dots \quad (18)$$

$$= \mathbf{p}^\mu \cdot \mathbf{v}^* - H^\mu - \min_i (\mathbf{p}_i \cdot \mathbf{v}^* - H_i) \quad (19)$$

$$+ (\mathbf{p}^\mu - \mathbf{p}^*) \cdot (\mathbf{v}^\mu - \mathbf{v}^*) + \dots \quad (20)$$

Lines (19) and (20) contain nonnegative quantities, which are positive if  $\mathbf{p}^\mu \neq \mathbf{p}^*$ . To see this for (20), observe that by the strict convexity  $H(t, \mathbf{x}, \mathbf{p}^\mu) > H(t, \mathbf{x}, \mathbf{p}^*) + (\mathbf{p}^\mu - \mathbf{p}^*) \cdot \mathbf{v}^*$  and  $H(t, \mathbf{x}, \mathbf{p}^*) > H(t, \mathbf{x}, \mathbf{p}^\mu) + (\mathbf{p}^* - \mathbf{p}^\mu) \cdot \mathbf{v}^\mu$ , and add these two inequalities.

On the other hand, the superdifferential of the function

$$\lim_{\mu \downarrow 0} \psi^\mu(t + \tau, \mathbf{x} + \boldsymbol{\xi}) = \phi(t + \tau, \mathbf{x} + \boldsymbol{\xi}) - \mathbf{p}^* \cdot \boldsymbol{\xi} + H^* \tau$$

at  $(\tau = 0, \boldsymbol{\xi} = 0)$  contains zero because  $\mathbf{p}^*$  corresponds to an admissible velocity. Therefore in the linear approximation the point  $\boldsymbol{\xi} = 0$  and other points of the shock surface of  $\psi$  tangent to  $(1, \mathbf{v}^*)$  are local maxima of  $\psi$  up to terms of higher order, and the full time derivative of  $\psi$  along the curve  $\boldsymbol{\xi} = \mathbf{v}^* \tau$  vanishes. As  $d\phi^\mu/dt$  is positive for trajectories that start outside the shock, we see that the flow (16) collapses them asymptotically on the shock surface. This completes the argument.

The full details of this proof and the rigorous derivation of uniqueness of the limit flow  $\bar{\gamma}$  will be given in a forthcoming article [17]. Here we just formulate the main result.

**Theorem 3** *Let  $\phi$  be a viscosity solution to the Cauchy problem for the Hamilton–Jacobi equation (1) with initial data (8). There exists a unique flow  $\gamma_{\mathbf{y}}$  of continuous trajectories such that  $\gamma_{\mathbf{y}}(0) = \mathbf{y}$ ,  $\dot{\gamma}_{\mathbf{y}}(t+0)$  is defined for all  $t$ ,  $\mathbf{y}$  and coincides with the admissible velocity  $\mathbf{v}^*(t, \gamma_{\mathbf{y}}(t))$  given by solution to the convex minimization problem for (13). The trajectory  $\gamma_{\mathbf{y}}$  continuously depends on  $\mathbf{y}$ . After a trajectory  $\gamma_{\mathbf{y}}$  comes to a shock, it stays inside the shock manifold for all later times. The flow is coalescing: if two trajectories  $\gamma_{\mathbf{y}'}$  and  $\gamma_{\mathbf{y}''}$  coincide at time  $t$ , then  $\gamma_{\mathbf{y}'}(t') = \gamma_{\mathbf{y}''}(t')$  for all  $t' > t$ .*

#### 4. Concluding remarks

Starting from a viscosity solution  $\phi$  to the Hamilton–Jacobi equation, we have constructed a unique continuous coalescing flow  $\gamma_{\mathbf{y}}$  compatible with the admissible velocity field  $\mathbf{v}^*$  defined in Section 2 in the sense that  $\dot{\gamma}_{\mathbf{y}}(t+0) = \mathbf{v}^*(t, \gamma_{\mathbf{y}})$ . This flow is a natural extension of the smooth flow defined by Hamilton’s equation (7). We conclude with a few observations concerning this construction.

1. Recall an important observation made in the original work of Bogachevsky: pieces of the shock manifold, irrespective of their dimension, are classified into *restraining* and *nonrestraining* depending on whether  $\mathbf{p}^*$  belongs to the interior or the boundary of the convex polytope formed by projection of  $\partial\phi(t, \mathbf{x})$  on the  $\mathbf{p}$  space. Particles stay on restraining shocks but leave non-restraining shocks along pieces of shock manifold of lower codimension corresponding to faces of the boundary containing  $\mathbf{p}^*$ . Shocks of codimension one are always restraining; in particular, so are all shocks in the one-dimensional case, which makes the construction of the coalescing flow  $\gamma_{\mathbf{y}}$  trivial, as remarked in the introduction. Interestingly, this classification of shocks, introduced in [12] (“acute” and “obtuse” superdifferentials of  $\phi$ ), seems to have been overlooked by physicists before.

2. Note that the construction of the admissible velocity  $\mathbf{v}^*$  is purely kinetic: when the Lagrangian is “natural,” i.e., has the form  $L(t, \mathbf{x}, \mathbf{v}) = K(\mathbf{v}) - U(t, \mathbf{x})$ , the value  $\mathbf{v}^*$  is the same for all choices of the potential term  $U(t, \mathbf{x})$  as long as kinetic energy  $K(\mathbf{v})$  is fixed. We owe this observation to P. Choquard.

3. Seen as a family of continuous maps of variational origin from initial coordinates  $\mathbf{y}$  to current coordinates  $\mathbf{x}$ , the flow  $\gamma_{\mathbf{y}}$  is clearly relevant for optimal transportation problems [18, 19]. An interesting problem suggested by B. Khesin is to study the extremal properties of this flow. Indeed it is known from [20] that before the first

shock formation the flow  $\gamma_y$  is and action minimizing flow of diffeomorphisms, while the first shock formation time  $t^*$  marks a conjugate point in the corresponding variational problem. According to the suggested view, the flow constructed above may be seen as a kind of saddle-point, rather than minimum, for a suitable transport optimization problem.

4. Another natural context to place our construction in is that of differential inclusions (see, e.g., [21]). The flow constructed here may be seen as a solution of differential inclusion

$$\dot{\gamma} \in \nabla_p H(t, \gamma, \text{Pr}_p \partial\phi(t, \gamma)), \quad (21)$$

where  $\text{Pr}_p$  is the  $p$  projection of the superdifferential  $\partial\phi$ . In comparison with standard constructions of the theory of differential inclusions the flow  $\gamma_y$  solves (21) in a stronger sense: the forward derivative  $\dot{\gamma}_y(t+0)$  exists everywhere. Also, a simple modification of Lemma 2 gives a proof of uniqueness for inclusion (21).

5. The flow  $\gamma_y$  was constructed as a limit of a parabolic regularization, and it was noticed above that the limit of  $(-\partial\phi^\mu/\partial t, \nabla\phi^\mu)$  belongs to  $\partial\phi(t, \mathbf{x})$  for any  $(t, \mathbf{x})$ . This statement can be refined if one considers the values of  $\nabla\phi^\mu$  along trajectories of the flow (16). Namely, arguments of Section 3 imply that under the successive limits  $\mu \downarrow 0$  and  $\tau \downarrow 0$ , the gradient  $(\partial\phi^\mu/\partial t, \nabla\phi^\mu)$  taken at  $(t+\tau, \gamma^\mu(t+\tau))$  converges to  $(-H^*, \mathbf{p}^*)$ . Therefore

$$\begin{aligned} \lim_{\tau \downarrow 0} \lim_{\mu \downarrow 0} \mu \nabla^2 \phi^\mu(t+\tau, \gamma^\mu(t+\tau)) &= \lim_{\tau \downarrow 0} \lim_{\mu \downarrow 0} \left[ \frac{\partial\phi^\mu}{\partial t} + H(t, \gamma^\mu(t+\tau), \nabla\phi^\mu) \right] \\ &= H(t, \mathbf{x}, \mathbf{p}^*) - H^* \\ &= \min_i (\mathbf{p}_i \cdot \mathbf{v}^* - H_i) - L(t, \mathbf{x}, \mathbf{v}^*) \\ &= \hat{L}(\mathbf{v}^*) = \min_{\mathbf{v}} \hat{L}(\mathbf{v}), \end{aligned}$$

where we took into account formulas (17) and (13). In other words, minimum of the convex minimization problem (13) coincides with the value of the “dissipative anomaly” in the parabolic regularization (4) of the Hamilton–Jacobi equation (1).

6. Observe also that convergence of superdifferentials makes it possible to use other smoothing procedures for  $\phi$  (e.g., convoluting it with a standard mollifier), giving the same limit  $\gamma_y$ . However, one can imagine the following completely different regularization of the discontinuous velocity field  $\nabla_p H(t, \mathbf{x}, \nabla\phi(t, \mathbf{x}))$ . Physically speaking, this regularization may be characterized by a “zero Prandtl number” in contrast with the previous class of regularizations featuring an “infinite Prandtl number.”

Consider the stochastic equation

$$d\gamma^\epsilon = \nabla_p H(t, \gamma^\epsilon, \nabla\phi(t, \gamma^\epsilon)) dt + \epsilon d\mathbf{W}(t),$$

where  $\mathbf{W}$  is the standard Wiener process. The corresponding stochastic flow is well defined in spite of the fact that  $\nabla\phi$  does not exist everywhere: whenever the trajectory  $\gamma^\epsilon$  hits shocks, noise in the second term will instantaneously steer it in a random direction away from them.

Assume that as  $\epsilon \downarrow 0$  the stochastic flow  $\gamma_y^\epsilon$  tends to a limit flow  $\tilde{\gamma}_y$ , which is also forward differentiable. It is easy to see that, due to the averaging, the forward velocity  $\mathbf{v}^\dagger(t, \tilde{\gamma}) := \dot{\tilde{\gamma}}_y(t+0)$  must belong to the convex hull of  $\mathbf{v}_j, j \in I(\mathbf{v}^\dagger)$ . Namely,

$$\mathbf{v}^\dagger = \sum_{j \in I(\mathbf{v}^\dagger)} \pi_j \mathbf{v}_j, \quad \pi_j \geq 0, \quad \sum_{j \in I(\mathbf{v}^\dagger)} \pi_j = 1, \quad (22)$$

where the velocities  $\mathbf{v}_j(t, \mathbf{x})$  are Legendre transforms of the corresponding momenta  $\mathbf{p}_j = \nabla \phi_j(t, \mathbf{x})$  at a singular point  $(t, \mathbf{x})$ . The coefficients  $\pi_i$  are equal to asymptotic values of shares of time that a trajectory  $\gamma^\epsilon$  spends in each of the domains where  $\phi = \phi_j$ . Condition (22) above is another compatibility condition, in certain sense dual to the admissibility condition in Section 2. For want of a better term let us call such a velocity  $\mathbf{v}^\dagger$  *self-consistent*.

The self-consistent velocity is a convex combination of *velocities* seen by an infinitesimal observer leaving  $(t, \mathbf{x})$  with velocity  $\mathbf{v}^\dagger$ . Compare this with the definition of admissible momentum  $\mathbf{p}^*$ , which is a convex combination of *momenta* seen by a similar observer moving with velocity  $\mathbf{v}^*$ . When  $H(t, \mathbf{x}, \mathbf{p}) = |\mathbf{p}|^2/2$  and  $\mathbf{v} = \mathbf{p}$ , self-consistent velocities and admissible velocities coincide. It is however clear that in the case of a general nonlinear Legendre transform  $\mathbf{v}^\dagger \neq \mathbf{v}^* = \nabla_{\mathbf{p}} H(t, \mathbf{x}, \mathbf{p}^*)$ .

In view of the analogy between self-consistent velocities and admissible momenta, it is tempting to conjecture that the admissible velocity is also unique. Generally speaking, this statement is wrong, although it holds in one and two space dimensions. In higher dimensions there exist Hamiltonians and sets of limiting momenta for which there is more than one admissible velocity [17]. It is an interesting problem nevertheless to see whether a limiting flow still exists in the limit of weak noise in spite of nonuniqueness of admissible velocity. This problem carries a certain similarity with the problem of limit behaviour for one-dimensional Gibbs measures in the zero-temperature limit, in the case of nonunique ground states.

## References

- [1] Bec J and Khanin K 2007 Burgers turbulence *Phys. Rep.* **447** 1–66
- [2] Kruzhkov S 1975 Weak solutions for Hamilton–Jacobi equations of the eikonal type *Mat. Sbornik* **98** 450–93 (in Russian)
- [3] Lions P-L 1982 *Generalized Solutions of Hamilton–Jacobi Equations (Research Notes in Mathematics vol 69)* (Boston: Pitman)
- [4] Crandall M G, Ishii H and Lions P-L 1992 User’s guide to viscosity solutions of second order partial differential equations *Bull. Amer. Math. Soc.* **27** 1–67
- [5] Hopf E 1950 The partial differential equation  $u_t + uu_x = \mu u_{xx}$  *Commun. Pure Appl. Math.* **3** 201–30
- [6] Lax P 1954 Weak solutions of nonlinear hyperbolic equations and their numerical computation *Comm. Pure Appl. Math.* **7** 159–93
- [7] Oleinik O 1954 On the Cauchy problem for nonlinear equations in the class of discontinuous functions *Dokl. Akad. Nauk SSSR* **95** 451–5 (in Russian)
- [8] Fleming W H and Soner H M 2005 *Controlled Markov Processes and Viscosity Solutions (Stochastic Modelling and Applied Probability vol 25)* (Berlin etc: Springer)
- [9] Fathi A 2009 *Weak KAM Theorem in Lagrangian Dynamics* (Cambridge: Cambridge University Press)

- [10] E W, Khanin K M, Mazel A E, Sinai Ya G 2000 Invariant measures for Burgers equation with stochastic forcing *Ann. of Math.* **151** 877–960
- [11] Dafermos C M 2005 *Hyperbolic Conservation Laws in Continuum Physics* (*Grundlehren in Mathematischen Wissenschaften* vol 325) (Berlin etc: Springer)
- [12] Bogaevsky I 2004 Matter evolution in Burgulence *Preprint* math-ph/0407073
- [13] Bogaevsky I 2006 Discontinuous gradient differential equations and trajectories in the calculus of variations *Mat. Sbornik* **197** 1723–51
- [14] Cannarsa P and Sinestrari C 2004 *Semiconcave Functions, Hamilton–Jacobi Equations, and Optimal Control* (*Progress in nonlinear differential equations and their applications* vol 58) (Boston: Birkhäuser)
- [15] Rockafellar R T 1970 *Convex Analysis* (Princeton: Princeton University Press)
- [16] Bregman L M 1967 The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming *USSR Computational Mathematics and Mathematical Physics* **7** 200–17
- [17] Khanin K and Sobolevski A 2009 *On dynamics of Lagrangian particles for Hamilton–Jacobi equations*, in preparation
- [18] Gangbo W and McCann R J 1996 The geometry of optimal transportation *Acta Math.* **177** 113–61
- [19] Villani C 2009 *Optimal Transport: Old and New* (Berlin etc: Springer)
- [20] Khesin B A and Misiołek G 2007 Shock waves for the Burgers equation and curvatures of diffeomorphism groups *Analysis and singularities. Part 2, Collected papers. Dedicated to Academician Vladimir Igorevich Arnold on the occasion of his 70th birthday* vol. 259 of *Tr. Mat. Inst. Steklova* (Moscow: Nauka) pp 77–85 (also available as preprint math/0702196)
- [21] Aubin J-P and Cellina A 1984 *Differential Inclusions: Set-Valued Maps and Viability Theory* (Berlin etc: Springer)